MASSACHUSETTS INSTITUTE of TECHNOLOGY
LIBRARIES

Digitized by the Internet Archive
in 2011 with funding from
Boston Library Consortium Member Libraries

http://www.archive.org/details/nashperfectequil00fude

HB31
.M415
no.499

MASS. INST. TECH.
AUG 23 1988
LIBRARIES

working paper
department
of economics

NASH AND PERFECT EQUILIBRIA OF DISCOUNTED REPEATED GAMES

By

Drew Fudenberg and Eric Maskin

No. 499 July 1988

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

NASH AND PERFECT EQUILIBRIA OF DISCOUNTED REPEATED GAMES

By

Drew Fudenberg and Eric Maskin

No. 499 July 1988

M.I.T. LIBRARIES
AUG 2 3 1966
RECEIVED

NASH AND PERFECT EQUILIBRIA OF DISCOUNTED REPEATED GAMES

By

D. Fudenberg* and E. Maskin **

February 1987

Revised July 1988

* Massachusetts Institute of Technology

** Harvard University and St. John's College, Cambridge

We thank two referees for helpful comments. Research support from the U.K. Social Science Research Council, National Science Foundation Grants SES-8607229 and SES-8520952 and the Sloan Foundation is gratefully Acknowledged.

ABSTRACT

The "perfect Folk Theorem" for discounted repeated games establishes that the sets of Nash and subgame-perfect equilibrium payoffs are equal in the limit as the discount factor δ tends to one. We provide conditions under which the two sets coincide before the limit is reached. That is, we show how to compute $\underline{\delta}$ such that the Nash and perfect equilibrium payoffs of the δ-discounted game are identical for all $\delta > \underline{\delta}$.

1. Introduction

The "Folk Theorem" for infinitely repeated games with discounting asserts that any feasible, individually rational payoffs (payoffs that strictly Pareto dominate the minmax point) can arise as Nash equilibria if the discount factor δ is sufficiently near one. Our [1986] paper established that under a "full-dimensionality" condition (requiring the interior of the feasible set to be nonempty) the same is true for perfect equilibria, so that in the limit as δ tends to 1 the requirement of subgame-perfection does not restrict the set of equilibrium payoffs. (Even in the limit, perfection does, of course, rule out some Nash equilibrium strategies.)

This paper shows that when the minmax point is in the interior of the feasible set and a second mild condition holds, the Nash and perfect equilibrium payoffs coincide before the limit. That is, for any repeated game, there is a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$, the Nash and perfect equilibrium payoffs of the δ-discounted game are identical. Our proof is constructive, and gives an easily computed expression for the value of $\underline{\delta}$. The payoff-equivalence result holds even though for any fixed $\delta < 1$ there will typically be individually rational payoffs that cannot arise as equilibria. In other words, the payoff sets coincide before attaining their limiting values. Payoff-equivalence is not a consequence of the Folk Theorem; indeed, it requires the additional conditions that we impose.

The key to our argument is the construction of "punishment equilibria", one for each player, that hold a player to exactly his reservation value. As in our other work on discounted repeated games ([1986], [1987a]), we interweave dynamic programming arguments and "geometrical" heuristics.

Section 2 introduces our notation for the repeated game model. Section 3 presents the main results. Section 4 provides counterexamples to show that the additional restrictions we impose are necessary, and that these

restrictions are not so strong as to imply that the Folk Theorem itself obtains for a fixed discount factor less than one. Through Section 4 we make free use of the possibility of public randomization. That is, we suppose that there exists some random variable (possibly devised by the players themselves) whose realizations are publicly observable. Players can thus use the random variable to (perfectly) coordinate their actions. In Section 5, however, we show that our results do not require public randomization.

2. Notation

We consider a finite n-player game in normal form

$$g: A_1 \times \ldots \times A_n \to R^n,$$

where $g(a_1,\ldots,a_n) = (g_1(a_1,\ldots,a_n),\ldots,g_n(a_1,\ldots,a_n))$ and $g_i(a_1,\ldots,a_n)$ is player i's payoff from the vector of actions $(a_1,\ldots,a_n)$. Player i's mixed strategies, i.e., the probability distributions over A_i, are denoted Σ_i. For notational convenience, we will write "$g(\sigma)$" for the expected payoffs corresponding to σ.

In the repeated version of g, each player i maximizes the normalized discounted sum π^i of his per-period payoffs, with common discount factor δ:

$$\pi^i = (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} g_i(\sigma(t)),$$

where $\sigma(t)$ is the vector of mixed strategies chosen in period t.

Player i's strategy in period t can depend on the past actions of all players, that is on the sequence $\{a(\tau)\}_{\tau<t}$, but not on the past choices of randomizing probabilities $\sigma(\tau)$. Each period's play can also depend on the realization of a publicly observed random variable such as sunspots. Although we feel that allowing public randomization is as reasonable as

prohibiting it, our results do not require it. Section 5 explains how, for δ close to 1, the effect of sunspots can be duplicated by deterministic cycles of play.

To give a formal definition of the strategy spaces, let $\omega(t)$ be a sequence of independent random variables with the uniform distribution on [0,1]. The history at time t is $h^t = (h^{t-1}, a(t-1), w(t))$, and a strategy for player i is a sequence s_i^t where s_i^t: $H^t \to \Sigma_i$ and H^t is the space of time t histories.

For each player j, choose minmax strategies $m^j = (m_1^j,\ldots,m_n^j)$, where

$$(0) \qquad m_{-j}^j \in \arg\min_{m_{-j}} \max_{m_j} g_j(m_j, m_{-j}) \text{ and } g_j(m^j) = \max_{a_j} g_j(a_j, m_{-j}^j).$$

Let

$$v_j^* = g_j(m^j),$$

where "m_{-j}" is a mixed strategy selection for players other than j, and $g_j(a_j, m_{-j}^j) = g_j(m_1^j,\ldots,m_{j-1}^j, a_j, m_{j+1}^j,\ldots,m_n^j)$.

We call v_j^* player j's reservation value. Since one feasible strategy for player j is to play in each period a static best response to that period's play of his opponents, player j's average payoff must be at least v_j^* in any equilibrium of g, whether or not g is repeated. Note that any Nash equilibrium path of the repeated game can be enforced by the threat that any deviation by j will be punished by the other players' minmaxing j (i.e., playing m_{-j}^j) for the remainder of the game.

Henceforth we shall normalize the payoffs of the game g so that $(v_1^*,\ldots v_n^*) = (0,\ldots,0)$. Call $(0,\ldots,0)$ the minmax point, and take $\bar{v}_i = \max_a g_i(a)$. Let

$$U = \{(v_1,\ldots,v_n) \mid \text{there exists } (a_1,\ldots,a_n) \in A_1 \times\ldots\times A_n \text{ with } g(a_1,\ldots,a_n) = (v_1,\ldots,v_n)\},$$

V = Convex Hull of U,

and

$$V^* = \{(v_1,\ldots,v_n) \in V \mid v_i > 0 \text{ for all } i\}.$$

The set V consists of feasible payoffs, and V^* consists of feasible payoffs that strictly Pareto dominate the minmax point. That is, V^* is the set of feasible, strictly individually rational payoffs.

3. Nash and Perfect Equilibrium

Any feasible vector of payoffs $(v_1,\ldots,v_n)$ that gives each player i at least $(1-\delta)\ \underline{v}_i$ is attainable in a Nash equilibrium, since Nash strategies can specify that any deviator from the actions sustaining $(v_1,\ldots,v_n)$ will be minmaxed forever. In a subgame-perfect equilibrium, however, the punishments must themselves be consistent with equilibrium play, so that the punishers must be given an incentive to carry out the prescribed punishments. One way to try to arrange this is to specify that players who fail to minmax an opponent will be minmaxed in turn. However, such strategies may fail to be perfect, because minmaxing an opponent may be more costly than being minmaxed oneself. Still, even in this case, one may be able, as in our [1986] paper, to induce players to minmax by providing "rewards" for doing so.

In fact, the present paper demonstrates that under certain conditions these rewards can be devised in such a way that the punished player is held to <u>exactly</u> her minmax level. When this is possible, the set of Nash and

perfect equilibrium payoffs coincide, as the following lemma asserts.

Lemma 1: For discount factor δ, suppose that, for each player i, there is a perfect equilibrium of the discounted repeated game in which player i's payoff is exactly zero. Then the sets of Nash and perfect equilibrium payoffs (for δ) coincide.

Proof: Fix a Nash equilibrium s^*, and construct a new strategy $\hat{s}$ that agrees with s^* along the equilibrium path, but specifies that, if player i is the first to deviate from s^*, play switches to the perfect equilibrium that holds player i's payoff to zero. (If several players deviate simultaneously, the deviations are ignored.) Since zero is the worst punishment that player i could have faced in s^*, he will not choose to deviate from the new strategy $\hat{s}$. By construction, $\hat{s}$ is a perfect equilibrium with the same payoffs as s^*.

Q.E.D.

Remark: Note that the lemma does not conclude that all Nash equilibrium strategies are perfect.

A trivial application of Lemma 1 is to a game, like the prisoners' dilemma, in which there is a one-shot equilibrium that gives all players their minmax values. An only slightly more complex case is a game where each player prefers to minmax than to be minmaxed, i.e., a game in which $g_i(m^j) > 0$ for all $i \neq j$. In such a game we need not reward punishers to ensure their compliance but can simply threaten them with future punishment if they fail to punish an opponent.

Theorem 1: Suppose that for all i and j, $i \neq j$, m_j^i, as defined by (0), is a pure strategy, and that $g_j(m^i) > 0$. Let $\underline{\delta}$ satisfy $\bar{v}_j(1-\delta) < \min_{i \neq j} g_j(m^j)$ for all i. Then for all $\delta \in (\underline{\delta}, 1)$, the sets of Nash and perfect equilibrium payoffs of the repeated game exactly coincide.

Proof: For each player i, define the i^{th} "punishment equilibrium" as follows. Players play according to m^i until some player $j \neq i$ deviates. If this occurs, they switch to the punishment equilibrium for j. Player i has no incentive to deviate from the i^{th} punishment equilibrium because in every period he is playing his one-shot best response. Player $j \neq i$ may have a short-run gain to deviating, but doing so results in his being punished, so that the maximum payoff to deviation is $\bar{v}_j(1-\delta)$,[1] which is less than $\min_{i \neq j} g_j(m^i)$ by assumption. So the hypotheses of Lemma 1 are satisfied. Q.E.D.

Remark 1: If the minmax strategies are mixed instead of pure, the construction above is inadequate because player j may not be indifferent among all actions in the support of m_j^i. Example 1 of Section 4 shows that in this case Theorem 1 need not hold.[2]

Remark 2: The proof of Theorem 1 actually shows that all feasible payoff vectors that give each player j at least $\min_{i \neq j} g_j(m_j^i)$ can be attained in

---

[1]Recall that we are expressing players' payoffs in the repeated game as discounted payoffs, and not as present values.

[2]However, in two-player games we can sharpen Theorem 1 by replacing its hypotheses with the condition that for all i and j, $i \neq j$, and all a_j in the support of m_j^i, $g_j(a_j, m_{-j}^i)$ is positive. Note that this condition reduces to that of Theorem 1 if all the m_j^i are pure strategies.

equilibrium if δ exceeds the $\underline{\delta}$ defined in the proof.

Although the hypotheses of Theorem 1 are not pathological (i.e., they are satisfied by an open set of payoffs in nontrivial normal forms), they do not apply to many games of interest. We now look for conditions that apply even when minmaxing an opponent gives a player less than her reservation utility.

In this case, to induce a player to punish an opponent we must give him a "reward" afterwards, as we explained earlier. To construct equilibria of this sort, it must be possible to reward one player without also rewarding the player he punishes. This requirement leads to the "full-dimensionality" requirement we introduced in our earlier paper: the dimensionality of V should equal the number of players. However full dimensionality is not sufficient for the stronger results of this paper, as we show in Section 4. We must strengthen it to require that the minmax point $(0,\ldots,0)$ is itself in the interior of V. Moreover, we need assume that each player i has an action $\hat{a}_i$ such that $g_i(\hat{a}_i, m^i_{-i}) < 0$, so that when minmaxed, a player has an action for which he gets a strictly negative payoff. (From our normalization, his maximum payoff when minmaxed is zero.)

Theorem 2: Assume that (i) the minmax point is in the interior of V, and that (ii) for each player i there exists $\hat{a}_i$ such that $g_i(\hat{a}_i, m^i_{-i}) < 0$. Then there exists $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$, the sets of Nash and perfect equilibrium payoffs of the repeated game exactly coincide.

Corollary: Under the conditions of Theorem 2, for $\delta > \underline{\delta}$, any feasible payoff vector v with $v_i \geq \underline{v}_i\ (1-\delta)$ for all i can be attained by a perfect equilibrium.

Remark: Hypothesis (ii) of the theorem is satisfied by generic payoffs in normal forms with three or more actions per player. The interiority condition (i) is, however, not generic: the minmax point can be outside V for an open set of payoffs. The condition is important because it ensures that in constructing equilibria to reward a punisher the ratio of his payoff to that of the deviator can be made as large as we like.

Proof of Theorem 2: We begin in part (A) with the case in which all the minmax strategies m_j^i are pure, or, equivalently, each player's choice of a mixed strategy is observable. Part (B) explains how to use the methods of our [1986] paper to extend the construction to the case where the minmax strategies are mixed.

(A) Assume that each m_j^i is a pure strategy. For each player i, choose an action $\hat{a}_i$ such that $g_i(\hat{a}_i, m_{-i}^i) = -x_i < 0$. For $j \neq i$, let $y_j^i = -g_j(\hat{a}_i, m_{-i}^i)$. The equilibrium strategies will have 2n "states", where n is the number of players. States 1 through n are the "punishment states", one for each player; states n + 1 to 2n are "reward states". In the punishment state i, the strategies are: Play $(\hat{a}_i, m_{-i}^i)$ today. If there are no deviations, switch to state n + i tomorrow with probability $p_i(\delta)$ (to be determined), and remain in state i with complementary probability $1 - p_i(\delta)$. If player j deviates, then switch to state j tomorrow. In reward state n + i, players play actions to yield payoffs $v^i = (v_1^i, \ldots, v_n^i)$, which are to be determined. If player j deviates in a reward state, switch to punishment state j.

Choose $v^i >> 0$ in V^* so that, for $i \neq j$, $x_i v_j^i - v_i^i y_j^i > 0$ (this is possible because $0 \in$ int V). Now set $p_i(\delta) = x_i/\delta v_i^i$, and choose $\bar{\delta} > x_i/(v_i^i + x_i)$, so that for $\delta > \bar{\delta}$, $p_i(\delta) < 1$. This choice of $p_i(\delta)$ sets player i's

payoff starting in state i equal to zero if she plays as specified. If player j does not deviate, his payoff starting in state i, which we denote w_j^i, solves the functional equation

$$(1) \quad w_j^i = (1-\delta)\,(-y_j^i) + \delta p_i(\delta) v_j^i + \delta(1-p_i(\delta)) w_j^i,$$

so that

$$(2) \quad w_j^i = (x_i v_j^i - y_j^i v_i^i) \,/\, (v_i^i + x_i).$$

By construction, the numerator in (2) is positive. The interiority condition has allowed us to choose the payoffs in the reward states so as to compensate the punishing players for punishing player i without raising i's own payoff above zero. Choose $\hat{\delta} < 1$ large enough that, for all i and j, $v_j^i > \bar{v}_j(1-\hat{\delta})$, and so that for $i \neq j$, $w_j^i > \bar{v}_j(1-\hat{\delta})$. Set $\underline{\delta} = \max(\hat{\delta}, \bar{\delta})$.

We claim that for all $\delta \in (\underline{\delta}, 1)$, the specified strategies are a perfect equilibrium. First consider punishment state i. In this state, player i receives payoff zero by not deviating. If player i deviates once and then conforms, she receives at most zero today (since she is being minmaxed) and she has a normalized payoff of zero from tomorrow on. Thus player i cannot gain by a single deviation, and the usual dynamic programming argument implies that she cannot gain by multiple deviations. Player j's payoff in state i is w_j^i (which exceeds $\bar{v}_j(1-\delta)$). A deviation could yield as much as $\bar{v}_j$ today, but will shift the state to state j, where j's payoff is zero, so player j cannot profit from deviating in state i. Finally, in reward state n+i, each player k obtains payoff v_k^i exceeding $\bar{v}_k(1-\delta)$, and so the threat of switching to punishment state k prevents deviations. The theorem now follows from Lemma 1.

(B) The strategies in part (A) punish player j if, in state i, he fails to use his minmax strategy m_j^i. The players can detect all deviation from m_j^i only if it is a pure strategy, or if the players' choices of mixed strategies are observable. (Otherwise, player j would not be detected if he deviated to a different mixed strategy with the same support). However, following our [1986] paper, we can readily modify the construction of (A) to allow for mixed minmax strategies. The idea is to specify that player j's payoff in reward state n+i depend on the last action he took in punishment state n in such a way that player i is exactly indifferent among all the actions in the support of m_j^i.

To begin, let $(a_j^i(k))$ be the pure strategies in the support of m_j^i, where the indexation is chosen so that

$$y_j^i(k) = -g_j(\hat{a}_i, a_j(k), m^i_{-j-i}) \leq -g_j(\hat{a}_i, a_j^i(k+1), m^i_{-j-i}) = y_j^i(k+1),$$

where m^i_{-j-i} is the vector of minmax strategies against i by players other than j. Thus $y_j^i(k)$ is player j's expected payoff in punishment state i if she plays her k^{th}-best strategy in the support of m_j^i. Next define

$$c = \max_{i, a_i, a_i^0, \sigma_{-i}} \| g_i(a_i, \sigma_{-i}) - g_i(a_i^0, \sigma_{-i}) \|.$$

This is the maximum variation a player's choice of action can induce in his own payoff. Also, let $\epsilon > 0$ be such that all payoff vectors v with $0 < v_i < 3\epsilon$ for all i are in the interior of V. (This is possible because $0 \in \text{int } V$.)

As in part (A), our strategies will have n punishment states and n reward states, with a probability $p_i(\delta)$ of switching from state i to state n+i if player i played $\hat{a}_i$ and each $j \neq i$ played an action in the support of m_j^i. However, when play switches to state n+i, player j's payoff depends on

the action that she took in the preceding period. Denote these payoffs by $v_j^i(k_j)$, where k_j is the index (as defined in the preceding paragraph) of the action last played by player j. Thus the vector of payoffs in state n+i is

$$v^i(k_1,\ldots,k_{i-1}, k_{i+1},\ldots,k_n) =$$

$$(v_1^i(k_1),\ldots,v_{i-1}^i(k_{i-1}), v_i^i, v_{i+1}^i(k_{i+1}),\ldots, v_n^i(k_n)).$$

This v^i is defined as follows: First choose v_i^i and $v_j^i(1)$ for each j to satisfy:

(3) $x_i v_j^i(1) - v_i^i y_j^i(1) > 0$, and

(4) $v_i^i < \epsilon x_i/c$, and

(5) $v_j^i(1) < \epsilon$.

These conditions can be satisfied because $0 \in$ int V. As in part (A), let $p_i(\delta) = (1-\delta)x_i/\delta v_i^i$. Now for each j and k_j, set

(6) $v_j^i(k_j) = v_j^i(1) + (1-\delta)\,[y_j^i(k_j) - y_j^i(1)] \,/\, \delta p_i(\delta)$.

With this specification of the reward payoffs, player j's payoff in state i is the same for each strategy in the support of m_j^i, and equals $[x_i v_j^i(1) - y_j^i(1)v_i^i] \,/\, (v_i^i + x_i)$, which is positive from inequality (3).

Now we must check that the specified payoffs for state n+i are all feasible, and that, for δ near one, no player wishes to deviate. Substituting the definition of $p(\delta)$ into (6), we have

$$(7) \quad v_j^i(k_j) = v_j^i(1) + [y_j^i(k_j) - y_j^i(1)]\,(v_i^i/x_i).$$

Referring to the bounds (4) and (5), we see that $v_j^i(k_j) < 3\epsilon$ for all j and k_j, and thus the vector v^i is feasible for all values of the k_j's. Finally, since player i's payoff in state i is zero, and his payoff in state n+j is bounded away from zero for all j, no player will wish to deviate in the reward states n+j. In state i, player $j \neq i$ obtains the same positive payoff from any strategy in the support of m_j^i, and she will be punished with zero for deviating from the support. Thus, for δ close to 1, player j will be willing to play m_j^i. The argument that player i will not deviate in state in i is exactly as in Case A. Q.E.D.

The interiority hypothesis of Theorem 2 implies that the set V has full dimension, i.e., that dim V = n. Let us briefly consider the connection between Nash and perfect equilibrium when V has lower dimension. When the number of players n exceeds two, our [1986] article shows by example that the Nash and perfect equilibrium payoff sets need not coincide even in the limit as δ tends to 1. Thus, for such examples, these sets do not coincide for $\delta < 1$. The story is quite different, however, for two player games.

Theorem 3: In a two-player game where dim V < 2, there exists $\underline{\delta} < 1$ such that, for all $\delta \in (\underline{\delta}, 1)$, the sets of Nash and perfect equilibrium payoffs coincide.

Proof: Let (m_1^2, m_2^1) be a pair of minmax strategies. (If there are multiple such pairs, choose one that maximizes player 1's payoff.) If $g(m_1^2, m_2^1) = (0,0)$, then (m_1^2, m_2^1) forms a Nash equilibrium of the one-shot game. In this

case, as infinite repetition of a one-shot Nash equilibrium constitutes a perfect equilibrium of the repeated game, application of Lemma 1 establishes the theorem. Suppose, therefore, that $g_i(m_1^2,m_2^1) < 0$ for some i. But then g cannot be a constant sum game and so, since dim V < 2, we can normalize the players' payoffs so that, for all (a_1,a_2), $g_1(a_1,a_2) = g_2(a_1,a_2)$. Take $\underline{v} = g_1(m_1^2,m_2^1)$ $(= g_2(m_1^2,m_2^1))$. Because $\underline{v} < 0$, there must exist (a_1^*,a_2^*) such that $g(a_1^*,a_2^*) = (v^*,v^*)$, where $v^* > 0$ (otherwise (m_1^1,m_2^1), where m_1^1 is a best response to m_2^1, is a minmax pair for which $g(m_1^1,m_2^1) = (0,0)$, contradicting the choice of (m_1^2,m_2^1)). We will show that, for δ near 1, there exists a perfect equilibrium of the repeated game in which both players' payoffs are zero. When m_1^2 and m_2^1 are pure strategies this is easily done by choosing $p(\delta)$ such that

$$(8) \quad (1-\delta)\underline{v} + \delta p(\delta) v^* = 0.$$

The equilibrium strategies consist of playing (m_1^2,m_2^1) in the first period and then either switching (permanently) to (a_1^*,a_2^*) with probability $p(\delta)$ or else starting again with probability $1-p(\delta)$. Deviations from this path are punished by restarting the equilibrium. Suppose that the support of m_1^2 is $\{a_1(1),\ldots,a(R)\}$ and that of m_2^1 is $\{a_2(1),\ldots,a_2(S)\}$. Suppose that the probability of $a_h(k)$ is $q_h(k)$. By definition,

$$(9) \quad \sum_j \sum_i q_1(i)\, q_2(j)\, g_1(a_1(i),a_2(j)) = \underline{v},$$

and since the m_j^i are minmax strategies,

$$(10) \quad \sum_i q_1(i)\, g_2(a_1(i),a_2(j)) \leq 0 \quad \text{for all } j$$

and

(11) $\sum_j q_2(j)\, g_1(a_1(i),a_2(j)) \leq 0$ for all i.

Now, in Lemma 2 below, we will show that, for all i and j, there exists $c_{ij} \geq 0$ such that

(12) $\sum_i q_1(i)\, [g_2(a_1(i),a_2(j)) + c_{ij}] = 0$

and

(13) $\sum_j q_2(j)\, [g_1(a_1(i),a_2(j)) + c_{ij}] = 0$

Take

$$p_{ij}(\delta) = \frac{c_{ij}(1-\delta)}{\delta\, v^*} .$$

Then, if players play (m_1^2, m_2^1) in the first period and switch (forever) to (a_1^*, a_2^*) with probability $p_{ij}(\delta)$ if the outcome is $(a_1(i),a_2(j))$, their expected payoffs are (0,0) (from (12) and (13)). Furthermore (12) and (13) imply that the players are indifferent among actions in the supports of their minmax strategies. Q.E.D.

Remark: An immediate corollary of Theorem 3 is that the Folk Theorem holds for two player games of less than full dimension even when mixed strategies are not observable. (This case was not treated in our [1986] paper.)

Lemma 2: Suppose that $B = (b_{ij})$ is an $R \times S$ matrix and that $p = (p(1),\ldots,p(R))$ and $q = (q(1),\ldots,q(S))$ are probability vectors such that

(14) $p \circ B \leq 0$ and $B \circ q \leq 0$.

Then there exists an RxS matrix $C = (c_{ij})$ such that $p \circ (B+C) = (B+C) \circ q = 0$ and $c_{ij} \geq 0$ for all i,j.

Proof: Consider a row $b_{i\cdot}$ such that $p \circ b_{i\cdot} < 0$. Now if, for all j. $b_{\cdot j} \circ q = 0$, then (14) implies that $p \circ B = 0$, a contradiction of the choice of $b_{i\cdot}$. Hence there exists j such that we can increase b_{ij} while (14) continues to hold. Indeed we can increase b_{ij} until either $p \circ b_{i\cdot} = 0$ or $b_{\cdot j} \circ q = 0$. Continuing by increasing other elements of B in the same way, we eventually obtain $p \circ B = 0$ and $B \circ q = 0$. Let C be the matrix of the increases that we make to B. Q.E.D.

4. Counterexamples

We turn next to a series of examples designed to explore the roles of the hypotheses of Theorems 1 and 2. Example 1 shows that Theorem 1 need not hold when the minmax strategies are mixed. Example 2 establishes that the hypotheses of Theorems 1 and 2 do not imply that all individually rational payoffs can be attained for some δ strictly less than one. Examples 3 and 4 show that both hypotheses (i) and (ii) of Theorem 2 are necessary. The detailed arguments can be found in our [1987b] working paper.

Example 1

Consider the game in Table 1

		Player 2 L	Player 2 R
Player 1	U	1, -1	-1, 2
	D	-1, -1	1, 0

Table 1

Note that player 1 is minmaxed if and only if player 2 mixes with equal probabilities between L and R. Thus, it is easy to see that the hypotheses of Theorem 1 are satisfied except for the assumption that 2's minmax strategy be pure. Notice that, for any δ, player 1's payoff must be positive in any perfect (indeed, in any Nash) equilibrium, because 1's payoff could be zero only were he minmaxed every period, in which case player 2's payoff would be negative. This already suggests that the set of perfect equilibrium payoffs will be a proper subset of the Nash payoffs, since a zero payoff can be used as a punishment in a Nash equilibrium, but the most severe punishment in a perfect equilibrium is positive.

To see that this suggestion is correct, let w^* be the infimum of player one's payoff over all perfect equilibria where player two's payoff is zero, and let e^n be a perfect equilibrium where player two's payoff is zero and player one's payoff w^n is near w^*. In the first period of e^n, player 1 must play D with probability one, or player 2 could obtain a positive payoff. Moreover, it can be shown that player 2 randomizes between L and R in the first period if δ is near enough to 1 (see our working paper). Let us modify e^n by slightly raising the probability that player 2 places on L. Player 2's payoff is unchanged by this alteration, since he is indifferent between L and

R. Player 1's payoff, however, declines, and for w^n sufficiently close to w^*, player one's payoff $\tilde{w}$ will now be less than w^*, so the payoff $(\tilde{w},0)$ is not supportable in a perfect equilibrium. Now further modify the strategies so that if player 1 deviates to U in the first period he is minmaxed forever. Since, in e^n, player 1 was deterred from deviating to U with a positive punishment, the threat of a punishment payoff of zero will deter player from deviating to U when the probability that player 2 plays L in the first period is slightly greater than in e^n. The modified e^n is, therefore, a Nash equilibrium whose payoffs cannot arise in a perfect equilibrium.

Example 2

Consider the game in Table 2. It satisfies the assumptions of both Theorems 1 and 2, so that for sufficiently large discount factors the Nash and perfect equilibrium payoffs coincide. However, as we shall see, this is not a consequence of the Folk Theorem, since, for such discount factors not all individually rational payoffs can arise in equilibrium.

	L	M	R
U	1, 0	2, -1	-2, -1
D	3, 0	1, 2	0, 1

Table 2

Now, the feasible point (3,0) is contained in the limit of the Nash equilibrium payoff sets as δ tends to 1. However, for any fixed discount factor $\delta < 1$, the Nash equilibrium payoffs are bounded away from (3,0). To see this observe that if, for $\delta < 1$, there existed a Nash equilibrium with

average payoffs (3,0), players would necessarily play (D,L) every period. But in any given period, player two could deviate to M and thereby obtain a positive payoff, a contradiction.

Example 3

Consider the two-player game in Table 3a, in which (0,0) lies on the boundary of V instead of the interior. This game satisfies hypothesis (ii) but not (i) of Theorem 2, and we shall see that the theorem fails.

	L	R
U	0, 0	-1, -4
D	2, 1	0, -5

Table 3a

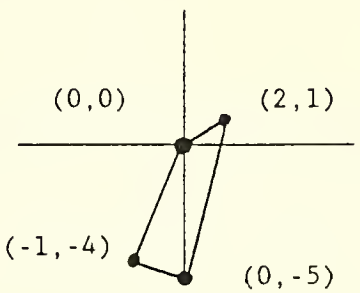

Figure 3b

As in Example 1, player 1's payoff cannot be exactly zero in a Nash equilibrium. (The only way that player 1's payoff could be zero along a path in which player 2's payoff is nonnegative would be for players to choose (U,L) with probability one in every period. But then player 1 could deviate to D in the first period and guarantee himself $2(1-\delta)$.) Indeed, as we show in our [1987b] paper, player one's Nash equilibrium payoff must be at least $14(1-\delta)/9$. Consider the perfect equilibrium where player 1's payoff is lowest.

If q is the probability that player 2 plays L in the first period, then player 1 can obtain at least $q(1-\delta) + \delta 14(1-\delta)/9$ by playing D in the first period. Playing R is costly for player 2, and so if q less than 1, player 2

must be rewarded in future periods. But, as Figure 3b shows, assigning player 2 a positive payoff induces an even higher payoff for 1. (This is because of the failure of the interiority condition.) The need to reward player 2 implies that player 1's equilibrium payoff exceeds $7(1-\delta)(1-q)$. Together the two constraints imply that it exceeds $7(1-\delta)/3$ but one can readily construct a Nash equilibrium where player 1's payoff is $2(1-\delta)$. Hence the Nash and perfect equilibrium payoffs do not coincide.

Example 4

Consider the game in Table 4.

	L	R
U	0, -1	-1, 0
D	0, -1	1, 1

Table 4

When player two plays L, player one obtains zero regardless of what he does, violating hypothesis (ii) of Theorem 2. Note, however, that hypothesis (i) (interiority) holds. Once again, player one's payoff exceeds zero in any Nash equilibrium. (Player 1 can be held to zero only if player 2 plays L every period, which is not individually rational.) From arguments similar to those for Example 1, we can show that for δ near 1, the payoffs in the Nash equilibrium that minimizes player 1's payoff among those where 2's is zero cannot be attained in any perfect equilibrium.

5. No Public Randomization

In the proof of Theorem 2, we constructed strategies in which play switches probabilistically from a "punishment" phase to a "reward" phase, with the switching probability chosen to make the punished player's payoff equal to zero. This switch is coordinated by a public randomizing device. The reward phase also relies on public randomization when the vector $v^i = (v_1^i,\ldots,v_n^i)$ lies outside the set U of payoffs attainable with pure strategies. Although public randomizing devices help simplify our arguments by convexifying the set of payoffs, they are not essential to Theorem 2. Convexification can alternatively be achieved with deterministic cycles over pure-strategy payoffs when the discount factor is close to one.

Our [1988] paper showed that public randomization is not needed for the proof of the perfect equilibrium Folk Theorem, even if players' mixed strategies are not observable. Lemma 2 of that paper established that, for all $\epsilon > 0$ there exist $\underline{\delta}$, such that for all vectors $v \in V^*$ with $v_i > \epsilon$ for i, and all $\delta > \underline{\delta}$, there is a sequence $\{a(t)\}_{t=1}^{\infty}$, where a(t) is a vector of actions in period t, whose corresponding normalized payoffs are v (i.e., $(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}g(a(t)) = v$) and for which the continuation payoffs at any time τ are with ϵ of v (i.e., for all τ, $\|(1-\delta)\sum_{t=\tau}^{\infty}\delta^{t-\tau}g(a(t))-v\| < \epsilon$. This result implies that, for δ large enough so that $\bar{v}_i(1-\delta) < \epsilon$, we can sustain the vector v^i as the payoffs of a perfect equilibrium where the equilibrium path is a deterministic sequence of action vectors whose continuation payoffs are always within ϵ of v^i, and where deviators are punished by assigning them a subsequent payoff of zero. Hence, attaining v^i, even when it does not belong to U, does not require public randomization.

Nor do we need public randomization to devise punishments whose payoffs re exactly zero. We can replace the punishment phase of Theorem 2 with one of deterministic length.

Proof of Theorem 2 without Public Randomization: (A) As in the earlier proof, we begin with the case of pure minmax strategies. Let y_j^i, x_i, and $\hat{a}_i$ be as before. Because $0 \in \text{int } V$ we can choose $\epsilon > 0$ and, for each i, a vector v^i with $v_j^i > 2\epsilon$ and $v_j^i x_i > y_1^i + 2\epsilon$ for all $j \neq i$. For δ near enough 1, we can choose function $v_i^i(\delta)$ so that

$$(8) \quad ||v_i^i(\delta) - v_i^i|| < \epsilon,$$

$$(9) \quad v_j^i x_i > y_j^i \, v_i^i(\delta) + \epsilon \text{ for all } j \neq i,$$

and such that there exists an integer $t(\delta)$ so that

$$(10) \quad (1-\delta^{t(\delta)})(-x_i) + \delta^{t(\delta)} v_i^i(\delta) = 0.$$

Let

$$(11) \quad w_j^i = (1-\delta^{t(\delta)})(-y_j^i) + \delta^{t(\delta)} v_j^i .$$

Substituting using (9) and (10), we have

$$(12) \quad w_j^i > \epsilon/x.$$

Take δ close enough to one so that for all i and j, $\bar{v}_j(1-\delta) < \min \{\epsilon, \epsilon/x_i\}$.

Now consider the following strategies: In state i, play $(\hat{a}_i, m_{-i}^i)$ for $t(\delta)$ periods, and then switch to state n+i, where play follows a deterministic sequence whose payoffs are $v^i(\delta) = (v_i^i(\delta), v_{-i}^i)$. For the play in state n+i we appeal to Lemma 2 of our [1988] paper, which guarantees that the continuation payoffs in state n+i are at least ϵ. If player i ever

deviates from his strategy, switch to the beginning of state i. By construction, player i's payoff is exactly zero at the beginning of state i, and increases thereafter, and since i is minmaxed in state i, he cannot gain by deviating. If player j≠i deviates in state i, he can obtain at most $\bar{v}_j(1-\delta) < \epsilon$, and from not deviating obtains at least w_j^i, which is larger. Finally, since payoffs at every date of each reward state are bounded below by ϵ, and deviations result in a future payoff of zero, no player will wish to deviate in the reward states.

(B) To deal with mixed minmax strategies, we must make player j's payoff in n+i depend on how he played in state i, as in the earlier proof of Theorem 2. We will be a bit sketchier here than before because the argument is essentially the same.

As before, $y_j^i(k)$ be player j's expected payoff from his k^{th} best action in the support of m_j^i, j≠i. Let

$$R_j^i = -(1-\delta)\sum_{\tau=0}^{\tau=t(\delta)-1}\delta^\tau[y_j^i(1)-y_j^i(k(\tau))];$$

this is the amount that player j sacrifices in state i relative to always playing $a_j^i(1)$. Now take

$$(13)\quad v_j^i(\delta) = v_j^i + R_j^i(1-\delta)/\delta^{t(\delta)}.$$

With these payoffs in state n+i, each player j is indifferent among all the actions in the support of m_j^i. If v_i^i and ϵ are taken small enough, $t(\delta)$ (as defined by (10)) will also be small, and so the right hand side of (13) will be feasible. Thus, the reward payoffs can be generated by deterministic sequences. Q.E.D.

References

Fudenberg, D. and E. Maskin [1986], "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information", Econometrica, 54, 533-554.

———— and ———— [1987a], "Discounted Repeated Games with Unobservable Actions, I: One-Sided Moral Hazard", mimeo.

———— and ———— [1987b], "Nash and Perfect Equilibria of Discounted Repeated Games", Harvard University Working Paper #1301.

———— and ———— [1988], "On the Dispensability of Public Randomization in Discounted Repeated Games", mimeo.

2610 026

3-22-89

6-5-09

Date Due

MAY 20 1990

MAR 0 1 1999

MAY 3 0 2000

Lib-26-67

MIT LIBRARIES
3 9080 005 243 834